Single-molecule studies of complex systems

Oijen, Antoine M. van

Published in:
Molecular BioSystems

DOI:
10.1039/b612545j

Publication date:
2007

Single-molecule studies of complex systems: the replisome

Antoine M. van Oijen

Received 30th August 2006, Accepted 15th November 2006
First published as an Advance Article on the web 22nd December 2006
DOI: 10.1039/b612545j

A complete, system-level understanding of biological processes requires comprehensive information on the kinetics and thermodynamics of the underlying biochemical reactions. A wide variety of structural, biochemical, and molecular biological techniques have led to a quantitative understanding of the molecular properties and mechanisms essential to the processes of life. Yet, the ensemble averaging inherent to these techniques limits us in understanding the dynamic behavior of the molecular participants. Recent advances in imaging and molecular manipulation techniques have made it possible to observe the activity of individual enzymes and record "molecular movies" that provide insight into their dynamics and reaction mechanisms. An important future goal is extending the applicability of single-molecule techniques to the study of larger, more complex multi-protein systems. In this review, the DNA replication machinery will be used as an example to illustrate recent progress in the development of various single-molecule techniques and its contribution to our understanding of the orchestration of multiple enzymatic processes in large biomolecular systems.

Introduction

Single-molecule techniques are rapidly changing the way we think about biochemical processes. A major strength of studying chemical reactions at the level of a single molecule instead lies in the direct measurement of distributions of molecular properties, rather than their ensemble averages. By constructing histograms of particular molecular observables for many individual molecules, deviant subpopulations can be identified and characterized. This static representation, however, does not allow for discrimination between a static heterogeneity and a dynamic process being responsible for the observed subpopulations. This point can be addressed by the recording of single-molecule trajectories, allowing us to follow biochemical processes in real time and observe any transient intermediate. The absence of a need for synchronization of the entire ensemble of molecules allows us to extract detailed dynamical information from single-molecule trajectories, otherwise obscured in kinetic ensemble studies by dephasing processes.

In recent years, a large number of biological systems have been the subject of study with single-molecule techniques. This development has already led to a greater understanding of the molecular mechanisms of biological processes which include, but are not limited to, active cellular transport,[1–6] muscle contraction,[7–11] ion transport,[12] ATP synthesis,[13–15] and redox reactions.[16] A field of significant activity within the single-molecule community is the study of processes that involve DNA. The robustness of DNA as a substrate and the development of techniques to manipulate individual DNA molecules (reviewed in (17)) contributed to exciting developments in the areas of transcription,[18–23] recombination,[24–30] repair,[31] and replication.[32–37]

An important future direction is the utilization of single-molecule techniques to unravel the orchestration of large macromolecular assemblies. Only a full understanding of the dynamics of proteins in the context of their physiologically relevant assemblies will allow for the incorporation of microscopic kinetic and thermodynamic information into a more global system-level description. This review will use the process of DNA replication to illustrate how single-molecule techniques can be employed to study large multi-protein assemblies.

DNA replication involves the coordinated activity of a large number of proteins. The replisome, the molecular machinery of DNA replication, unwinds the double-stranded DNA, provides primers to initiate synthesis, and polymerizes

Harvard Medical School, Dept. of Biological Chemistry and Molecular Pharmacology, 240 Longwood Avenue, Boston, MA 02115, US. E-mail: antoine_van_oijen@hms.harvard.edu; Fax: +1 617-738-0516; Tel: +1 617-432-5586

Antoine M. van Oijen

Antoine van Oijen received a MSc and PhD in physics from Leiden University, the Netherlands. He held a post-doctoral position at the Department of Chemistry and Chemical Biology at Harvard University. Currently, he is an assistant professor at the Department of Biological Chemistry and Molecular Pharmacology at Harvard Medical School. His research interests are centered on developing and using single-molecule techniques to study the activity of DNA-processing enzymes. He has been a recipient of a Searle Scholarship and a Career Award from the National Science Foundation.

nucleotides onto each of the two growing strands.[38,39] DNA is replicated at a truly impressive speed and accuracy: in *Escherichia coli* (*E. coli*), the replication fork moves at a rate approaching 1000 nucleotides per second[40] while making less than one mistake per 10^9 nucleotide incorporations.[41] Various specific interactions between the proteins at the replication fork give the replisome its efficiency. Remarkable progress has been made in characterizing the structural and functional properties of the individual components; their coordination at the replication fork is less well understood. The large number of tightly coupled chemical and mechanical activities involved in DNA replication make it an ideal system to illustrate the variety in observables that can be accessed by single-molecule techniques and the wealth of kinetic parameters that can be obtained from them. This review aims to describe the significant contributions made by single-molecule studies to our understanding of the functioning of the protein complexes involved in replication. First, the various methods to mechanically manipulate individual DNA molecules and study proteins interacting with them will be reviewed. Subsequently, single-molecule studies on the individual components of the replication machinery will be discussed. This review will conclude with a description of the progress made towards obtaining a single-molecule view of the entire replisome.

Methods to study nucleic acid enzymes at the single-molecule level

Mechanical manipulation of individual DNA molecules

The recent development of methods to mechanically manipulate single DNA molecules has led to a detailed understanding of the mechanical properties of DNA. The controlled stretching and twisting of individual DNA molecules allowed researchers to quantitatively describe the mechanical properties of double-stranded DNA in terms of theoretical models originally developed to describe the behavior of ideal polymers.[42,43] More recently, these techniques have opened the door for studying the activity of nucleic acid-processing enzymes at the single-molecule level. A brief overview of the methods used to manipulate DNA and to monitor the activity of nucleic acid enzymes is presented here.

Flow stretching. A laminar flow of aqueous buffer can be used to exert forces on objects by transfer of momentum from the fluid to the object. DNA can be attached on one end to the surface of a flow chamber and the other end to a polystyrene bead whose viscous drag exerts a force on the DNA and stretches it (Fig. 1a).[35,44] Alternatively, the drag on the DNA molecule itself can be utilized to stretch it and use it as a template for single-molecule studies.[31] Combining the flow-stretching with wide-field optical microscopy allows for a simultaneous observation of multiple reactions. This multiplexing is essential for gathering statistically significant sample sizes and for observing the low-probability events associated with the activity of complex, multi-protein machineries.

Magnetic tweezers. Magnetic fields can be used to manipulate and apply forces to DNA or proteins that are tethered to magnetic particles. The presence of a preferred magnetization axis in magnetic beads allows them to be aligned with the direction of the applied magnetic field.[45–47] This allows individual DNA molecules, bound to a surface on one end and to a magnetic bead to the other, to be over- or underwound, and thus supercoiling to be introduced (Fig. 1b). This method has proven to be particularly powerful in the study of DNA topoisomerases and gyrases.[46,48]

Optical trapping. Optical tweezers, or optical traps, take advantage of the fact that light exerts force on matter. Dielectric particles, such as polystyrene beads or bacteria, are attracted to the center of a tightly focused laser beam and can be trapped there.[49,50] The force exerted on the object depends on the power of the laser, the dimensions of the object, and the difference in index of refraction between the object and the surrounding medium. Forces can be exerted on DNA or on a protein bound to DNA by optically trapping a polystyrene bead that is attached to the free end of a DNA whose other end is either bound directly to a surface or to a surface-immobilized protein (Fig. 1c). The large range in forces that can be applied (100 fN–100 pN) makes optical trapping techniques suitable for the investigation of the effect of force on biochemical processes.[51] The high spatial accuracy associated with optical trapping has enabled measurements of translocating nucleic acid enzymes with a resolution of a single

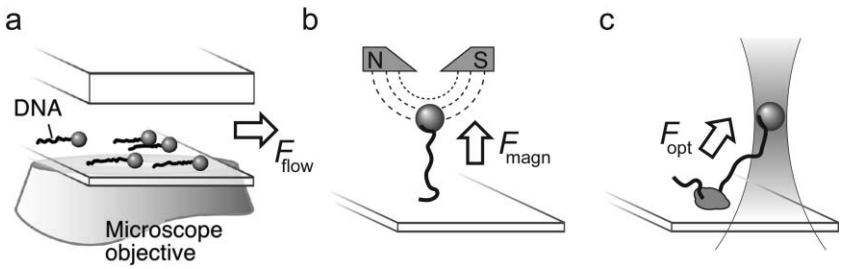

Fig. 1 Schematic depiction of methods to mechanically stretch individual DNA molecules. (a) Flow stretching. The Stokes drag force exerted by a laminar flow on a bead at the free end of surface-attached DNA will cause the molecules to be stretched close to and parallel to the surface of a microscope cover slip. (b) Magnetic tweezers. The bead attached to the free end of the DNA is paramagnetic and can be manipulated by a set of magnets directly above the sample. The presence of a magnetic dipole moment in the bead allows torque to be introduced into the DNA. (c) Optical trapping. A tightly focused laser beam traps the beam attached to the free end of the DNA, allowing the DNA to be stretched. In the case illustrated here, the DNA is held to the surface by a surface-immobilized nucleic acid-processing enzyme.

base pair.[19] Generally, only a small number of beads can be trapped simultaneously, making multiplexed observations difficult.

For all three methods, the absolute force exerted on the DNA can be obtained by two independent methods. First, the force-dependent DNA length can be fitted to the worm-like chain model, an analytical model that describes the DNA as a flexible rod that displays smooth, random bends as a result of thermal fluctuations. With only a single parameter, the bending persistence length, this model accurately predicts the DNA length as a function of force.[42] Second, the amplitude of the Brownian motion of the DNA-attached bead perpendicular to the direction of the force is directly related to the length of the DNA and the force applied.[45]

Fluorescence detection of individual DNA-binding proteins

Advances in fluorescence spectroscopy and microscopy have made it possible to detect the fluorescence from a single chromophore under biologically relevant conditions.[52] By covalently attaching fluorophores to proteins, their movement can be monitored by single-molecule imaging. Furthermore, by making use of spectroscopic properties and phenomenon such as polarization and energy-transfer, the protein's conformational dynamics can be followed at the microscopic level. Imaging the fluorescence of an individual labeled protein with a sufficiently large signal-to-background ratio requires an efficient suppression of fluorescence background from bulk solution. Total internal reflection fluorescence (TIRF) microscopy enables the selective illumination and excitation of fluorophores in a thin layer immediately adjacent to a glass-water interface and thus minimizes the volume of solution contributing to background (Fig. 2a).[53] The basic concept of TIRF microscopy is simple, requiring only an excitation light beam traveling at a high incident angle through a glass coverslip. At a specific critical angle of incidence, the beam of light is totally reflected from the glass/water interface, rather than passing through. The reflection generates a very thin, evanescent illumination (with a thickness of 100–200 nanometers) in the aqueous medium (Fig. 2a). Combining wide-field fluorescence imaging with total internal reflection illumination allows for a highly multiplexed observation of fluorescently labeled proteins interacting with surface-immobilized DNA molecules. Flow-stretching long DNA close to and parallel to the surface can be used with TIRF to allow for observation of the movement of fluorescently labeled nucleic acid enzymes along DNA.[31]

Spectroscopic properties of fluorescent labels can be used to gain information about the conformation of proteins and distances between them. Distance changes on a length scale comparable with the dimensions of biological macromolecules can be measured by fluorescence resonance energy transfer (FRET), a process where the excitation energy of a donor chromophore is transferred to an acceptor dye *via* an induced dipole–dipole interaction. The strong distance dependence of this interaction allows FRET to be used as a molecular ruler to probe small changes in distance between two proteins or between two sites on a protein.[54,55] Single-molecule FRET has been used to monitor conformational changes in surface-immobilized DNA,[56] in proteins on DNA,[57] and to probe movement of protein on DNA (Fig. 2b).[33]

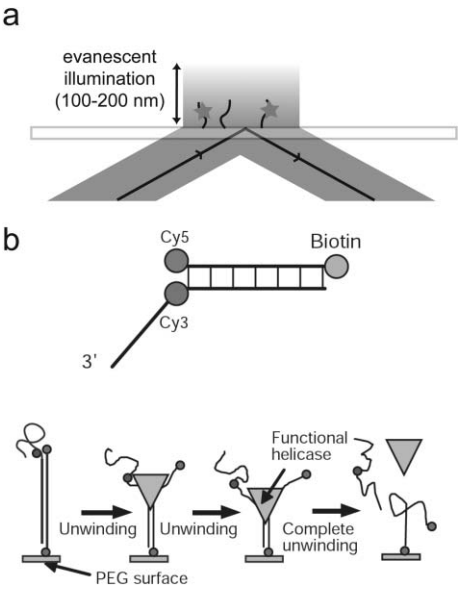

Fig. 2 (a) Total internal reflection microscopy (TIRF) restricts illumination to a thin layer immediately above the glass surface, thereby suppressing fluorescence background from bulk solution. The evanescent field is created by the total internal reflection of a collimated laser beam illuminating the sample at an oblique angle. (b) Schematic depiction of single-molecule Fluorescence Resonance Energy Transfer (FRET) experiment to probe helicase-mediated unwinding of double-stranded DNA. The transfer of excitation energy from the donor fluorophore (here Cy3) to the acceptor fluorophore (here Cy5) is extremely sensitive to the physical distance between the two probes. Unwinding of the duplex will cause this distance to increase allowing for detailed measurements of unwinding with sub-nanometer resolution (Figure adapted from ref. 70).

Synthesizing DNA: the DNA polymerase

DNA polymerases catalyze the synthesis of a new DNA strand on a ssDNA template.[58] Both prokaryotic and eukaryotic cells contain multiple DNA polymerase activities, involved in DNA replication, recombination, and repair. DNA polymerases extend a DNA chain by adding to it complementary nucleotides one at a time in such a way that the 3′-OH end of the growing chain attacks the alpha phosphorous atom of the incoming nucleotide's 5′-phosphate group in a S_N2 reaction. Most known DNA polymerases perform this reaction in a processive manner, incorporating thousands of nucleotides without dissociating from its template. All of the bacterial DNA polymerases possess a 3′–5′ exonucleolytic activity that proceeds in the reverse direction from DNA synthesis. This activity provides the capability of proofreading by removing a mismatched penultimate nucleotide.

Several kinetic and structural studies have demonstrated that the rate, processivity, and efficiency of proofreading all strongly depend on subtle conformational variations within the enzyme's active site.[59,60] The ability to observe a single DNA polymerase moving along DNA while incorporating nucleotides allows for a more extensive study of the relation

between enzyme-template interaction and kinetics. The rate of a single DNA polymerase synthesizing DNA on a single-strand DNA template can be measured by utilizing the difference between the elastic properties of single-stranded and double-stranded DNA.[36,37] Conversions between dsDNA and ssDNA can be monitored through a change in total DNA length, where the number of nucleotides converted is obtained by using the difference between the lengths of ssDNA and dsDNA (Fig. 3). The catalytic rates of the single enzymes were demonstrated to display large heterogeneity within the population of molecules. Furthermore, the enzymatic rate of a single enzyme fluctuated over time.[36] This observation is confirmed by single-molecule studies on other enzymatic systems[16,44] and is interpreted as small conformational changes in the enzyme subtly modulating the geometry of the active site.[61]

By means of optical[37] or magnetic tweezers,[36] the force can be changed in a controllable way and its influence on the polymerization rate investigated. For several prokaryotic polymerases, it was demonstrated that forces higher than 30–40 pN stall the enzyme.[36,37] Single-molecule studies on the DNA polymerase of the T7 bacteriophage showed that even higher forces stimulate the exonuclease activity by several orders of magnitude.[37] The response of the enzyme's translocation rates on a variation of the load applied provides information on which biochemical steps in the pathway are coupled to movement. The sensitivity of the polymerization rate to tension indicates that the rate-limiting step is directly affected by force. This behavior can be explained by a change in the conformation of the enzyme during the rate-limiting step.[59] A quantitative interpretation of the force-dependence observed in the single-molecule studies seemed to suggest that the DNA polymerases convert two or more single-stranded template bases to double-stranded DNA geometry during each catalytic cycle. Structural data suggests that only a single base is converted each cycle.[60] This discrepancy was resolved by modeling studies taking into account local interactions of single-stranded bases with the enzyme.[62] Every catalytic cycle, the interactions of the enzyme with the single-stranded DNA cause the DNA to shorten by an amount equivalent to n bases (where n is 2 or larger, depending on the polymerase). After condensation of one nucleotide into the DNA chain, a conformational change in the enzyme–substrate complex causes the DNA to lengthen again by $n - 1$ bases.[62]

The ability to precisely tune the polymerization and exonuclease rate of a single DNA polymerase allowed for a complete characterization of all the transition rates involved in both reactions. In combination with previously measured bulk-phase data, a theoretical description of the different states, and the transitions between them, of the DNA polymerase as a biochemical network could be obtained (Fig. 4).[63] Such biochemical networks allow the examination of the effects of control variables (such as tension, substrate concentration) on enzyme action and serve as building blocks of a systems approach to understanding the activity of these enzymes in the context of larger machineries.

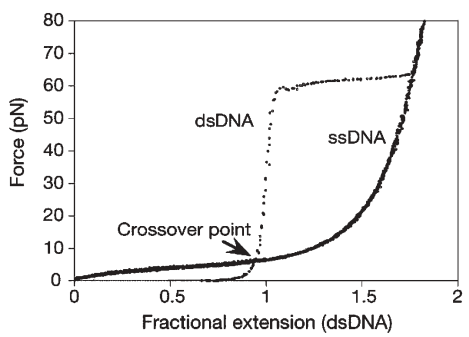

Fig. 3 Force-extension data of ssDNA and dsDNA. At forces below the crossover point (~6 pN), single-stranded DNA is shorter than double-stranded DNA. At higher forces, single-stranded become longer than the duplex. Enzymatically catalyzed conversions between the two states can be visualized by changes in DNA length (Figure adapted from ref. 37).

Unwinding double-stranded DNA: the helicase

DNA helicases are enzymes capable of unwinding duplex DNA to provide the single-stranded DNA templates that are required in many biological processes. All helicases separate the strands of a double helix using the energy derived from nucleotide hydrolysis. Single-molecule techniques are well suited to monitor rates and processivities of unwinding, and study their dependence on the applied load. There are five classes ('superfamilies') of helicases based on sequence homology.[64] Depending on the superfamily, helicases display different polarities of movement on single-stranded DNA and adopt different oligomerization states.

The unwinding activity of the superfamily I protein UvrD, a prokaryotic helicase involved in a number of repair pathways,[65] has been studied at the single-molecule level by using the fact that single-stranded DNA is longer than double-stranded DNA at high stretching forces (Fig. 3).[66] The helicase was observed to switch strands during unwinding, leading to a reversed translocation of the protein away from the fork and a gradual re-annealing of the two single-stranded DNA products. A Fourier analysis of the noise associated with the measurement of the extension of the DNA allowed the determination of the enzyme's step size.

Another illustrative example of a helicase that has been intensively investigated by different single-molecule techniques is RecBCD, an enzyme with both helicase and nuclease

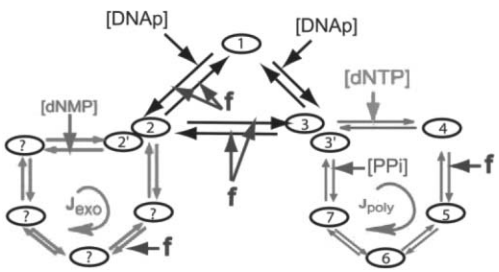

Fig. 4 Kinetic network of T7 DNA polymerase activity. Distinct cycles for polymerase (red) and exonuclease (green) pathways are linked by a cycle (black) involving binding or unbinding of DNA to the poly or exo active sites. Results from force-dependent single-molecule experiments led to the identification of the transitions that are sensitive to tension applied to the DNA (Figure from ref. 63, copyright by the National Academy of the United States).

activities that plays an important role in the repair of chromosomal DNA through homologous recombination.[65] Unwinding and nucleolytic activity was studied by monitoring the displacement of intercalating fluorescent dye from a flow-stretched DNA substrate.[30] Both this experiment and a previous study[67] pointed to a wide distribution of enzymatic rates. Encounter of the enzyme with the χ sequence, an 8 base sequence that regulates recombination in *E. Coli*,[68] was shown to change its nuclease activity.[30] Using optical tweezers to apply tension to a DNA template that was threaded through a surface-immobilized RecBCD enzyme, Perkins *et al*. observed DNA unwinding with a precision of 2 nm, implying a unitary step size of the enzyme below six base pairs.[69] Interestingly, they observed that the velocity of the enzyme remained constant for hundreds to thousands of base pairs, before suddenly switching to a different rate, pausing, or even occasionally sliding back along the DNA template.

Ha and coworkers pioneered the use of single-molecule FRET in the study of the molecular mechanisms of helicase activity. Association of *E. Coli* Rep helicase, also a superfamily I member, with a duplex DNA substrate and subsequent unwinding could be observed by a strategic labeling of the DNA with the donor molecule on one strand and the acceptor dye on the other (Fig. 2b).[70] Later experiments involved fluorescent labeling of the protein itself and led to the determination of the orientation of a Rep monomer bound to a single-stranded-double-stranded DNA, as well as the relative orientation of one of its sub-domains.[71] Furthermore, the authors demonstrated that an interaction between the helicase and the 3′ terminus of one of the unwound DNA strands prevents the protein from dissociating from the DNA.[33]

Many of the helicases that unwind DNA in the context of the replication machinery have not yet been studied at the single-molecule level. Well-known examples of prokaryotic replicative helicases are the *E.coli* DnaB, the bacteriophage T7 gp4, and T4 gp41, all hexameric, donut-shaped proteins belonging to the DnaB superfamily. Extensive biochemical and structural characterizations resulted in the picture that these proteins encircles a single strand of DNA and are able to translocate from the 5′ to the 3′ direction.[72] Upon encountering a single-stranded DNA/double-stranded DNA junction, the complementary strand is displaced and the double-stranded DNA unwound. It is not understood how the single-stranded DNA translocation activity is coupled to the unwinding, however. Two different models are typically considered: an active or passive coupling.[73] A passive helicase acts as a Brownian ratchet: it waits for the thermal fluctuation that transiently melts the first few base pairs of the double-stranded DNA, and then moves forward and binds to the newly available single-stranded DNA. The active model describes a helicase that employs an irreversible powerstroke to disrupt the double-stranded DNA. In this case, the hydrolysis of nucleotides is tightly coupled to the destabilization of the duplex, leading the helicase to unwind the double-stranded DNA without being significantly slowed down by the single-stranded DNA/double-stranded DNA barrier. A kinetic characterization at the single-molecule level of both unwinding of double-stranded DNA and translocation on single-stranded DNA will undoubtedly answer this and many other outstanding questions.

Single-stranded DNA-binding proteins

Single-stranded DNA-binding proteins (SSBs) are essential co-factors in a large number of processes involving DNA, such as replication[58] and recombination.[74] Not only do they coat the single-stranded DNA that is transiently exposed at the lagging strand during replication to protect it from nucleolytic degradation, they are also thought to play an important regulatory role within the replisome through interactions with other replication proteins. The development of various techniques to stretch individual DNA with a large range of forces allowed for the characterization of the thermodynamic and kinetic properties of single-stranded DNA coating and double-stranded DNA destabilization by SSBs.

An instructive example is the work by the Williams group that studied the effect of the T4 single-stranded DNA binding protein, the gene 32 protein, on the melting behavior of duplex DNA.[34,75] Thermodynamically analogous to the melting temperature of double-stranded DNA, a critical force exists at which a duplex DNA is converted into two single-stranded DNA molecules. The influence of SSBs on this melting force provides information of the binding affinity of the protein to single-stranded DNA. In contrast to temperature-induced melting, these force studies can be performed at physiological temperatures, thus avoiding protein denaturation. The dependence of the gp32 equilibrium binding constant as a function of the ionic strength revealed that protein binding is regulated by intramolecular conformational changes.

The replisome: A multi-enzyme replication machinery

Replisome assembly

Replication forks are assembled at specific chromosomal sites known as origins of replication. Before an active replisome is ready to unwind and synthesize DNA, the individual replication proteins need to be assembled onto the origin. The T4 bacteriophage replisome is a replication model system whose assembly pathways are intensively studied. Both the T4 DNA polymerase sliding clamp complex (homologous to the β subunit of the *E. coli* DNA polymerase III) and the T4 helicase form rings around DNA during replication and require accessory proteins to assemble.[76]

The assembly process is a pathway that contains many intermediates, a wide range of transition rates, and multiple branches. The order in which various proteins assemble is difficult, if not impossible, to monitor using ensemble-averaged experiments. The groups of Benkovic and Hammes have used single-molecule FRET between labeled T4 replication proteins to unravel the order of their assembly.[32,77–79] A low concentration of forked DNA substrates was immobilized on a surface and the fluorescence from the labeled replication proteins imaged while assembling. This work demonstrated that the T4 DNA polymerase can be assembled through one of four major different pathways.[77] These different routes of assembly each may play a role in different phases of the

replication cycle, such as leading- and lagging-strand synthesis. Using these single-molecule FRET imaging techniques, the authors also uncovered a mechanism that inhibits DNA polymerase activity until the helicase loading protein has completed its job of assembling the helicase.[78,79] These experiments characterized equilibrium binding stoichiometries by taking "snapshots" under various assembly conditions. An exciting future direction will be the real-time observation of the assembly process allowing for the determination of kinetic parameters.

Fork elongation

Once properly assembled, replication fork elongation commences and DNA is duplicated at a high rate and accuracy. The replication of parental DNA into two identical daughter molecules is a key example of how multiprotein machineries perform highly complex mechanical tasks at the microscopic level. The helicase unwinds the parental double-stranded DNA (dsDNA) into two DNA strands allowing two DNA polymerases, complexed with processivity factors, to each synthesize DNA on the single-stranded templates. The 5′ to 3′ direction of polymerase-dependent nucleic acid synthesis permits one of these enzymes to synthesize DNA in a continuous fashion on the leading strand, but forces the polymerase on the lagging strand to restart at short intervals, using short RNA primers made by a DNA primase. The discontinuous synthesis of DNA on the lagging strand gives rise to a succession of Okazaki fragments that are later processed and ligated into one continuous strand. Single-stranded DNA-binding proteins remove any secondary structure that may inhibit synthesis and protect the stretches of transiently exposed ssDNA from nucleolytic attacks.

The various enzymatic events involved in Okazaki fragment synthesis must occur with such organization that lagging-strand synthesis remains in step with the continuous leading-strand synthesis. This synchronization requires a precisely timed series of enzymatic steps that control the synthesis of a primer, the recycling of the lagging-strand DNA polymerase, and the production of an Okazaki fragment. Different models have been put forward to explain how these slow enzymatic steps can take place at the lagging strand without losing coordination with the continuous and rapid leading-strand synthesis.[80–82]

Lee *et al.* used single-molecule techniques to probe the kinetics of leading- and lagging-strand synthesis mediated by the replication machinery of bacteriophage T7.[35] Its replisome can be reconstituted *in vitro* with a small number of purified proteins (Fig. 5), making it an attractive model to study the orchestration at the replication fork. Nonetheless, the organization of the T7 replication system closely mimics that of *E. coli* and more complex organisms.[39] The T7 DNA polymerase consists of a 1 : 1 complex of the T7 gene 5 protein (gp5), encoded by the phage, and the thioredoxin processivity factor, encoded by the *E. coli* host.[83] The T7 gene 4 protein (gp4) assembles as a hexamer and provides both helicase and primase activities.[84,85] The helicase activity, required for unwinding the parental DNA strands, is located in the C-terminal half, and the primase activity, capable of

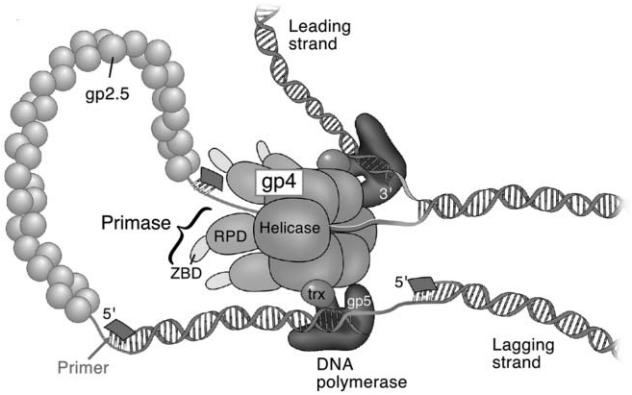

Fig. 5 The bacteriophage T7 replication machinery. The replisome consists of the hexameric T7 gene 4 protein (gp4) and two copies of the T7 DNA polymerase (T7 gene 5 protein (gp5) complexed with *E. coli* thioredoxin (trx)). T7 gene 2.5 protein (gp2.5), the ssDNA-binding protein, coats the transiently exposed ssDNA in the replication loop. Gp4 consists of a primase and a helicase domain, connected by a linker region. The primase domain consists of two subdomains: a zinc-binding domain (ZBD) and the RNA polymerase domain (RPD).

synthesizing the tetraribonucleotide primers that are required to initiate lagging-strand DNA synthesis, is located in the N-terminal half.

The authors stretched individual DNA molecules by laminar flow and monitored their lengths by tracking the positions of small beads attached to the ends of the DNA molecules. Conversion from double- to single-stranded DNA was monitored through a decrease in total DNA length at the low force used (Fig. 3).[44] By using a forked DNA template, a complex of one T7 DNA polymerase and helicase-primase could be assembled on one end of the DNA molecule. Leading-strand synthesis catalysed by T7 DNA polymerase converts one DNA strand arising from gp4 helicase activity into double-stranded DNA. In the absence of the lagging-strand DNA polymerase, the lagging strand will remain in the single-stranded form. By attaching the DNA to the surface of the flow cell by the 5′-end of the lagging strand, leading-strand synthesis could be detected by an effective shortening of the DNA (Fig. 6a).

A typical leading-strand synthesis trajectory mediated by the T7 DNA polymerase in association with the gp4 helicase is depicted in Fig. 6a (trace 1). Before the reaction was initiated, the flow cell was stringently washed with only buffer and nucleotides, thus effectively reducing the free protein concentration in the reaction volume to zero. As a result, protein exchange between the solution and the DNA-bound complex during the reaction was prevented. This situation, impossible to achieve using bulk biochemical techniques, allowed for a determination of the true processivity and rate of the replication complex.

The primase activity of the gp4 was studied by adding ribonucleotides to the reaction mixture. In the presence of ribonucleotides, short pauses occurred in the single-molecule leading-strand synthesis traces (Fig. 6a, trace 2; pauses are indicated by grey arrows). These pauses, with an average duration of several seconds, were demonstrated to result from primer synthesis by repeating these experiments with a gp4

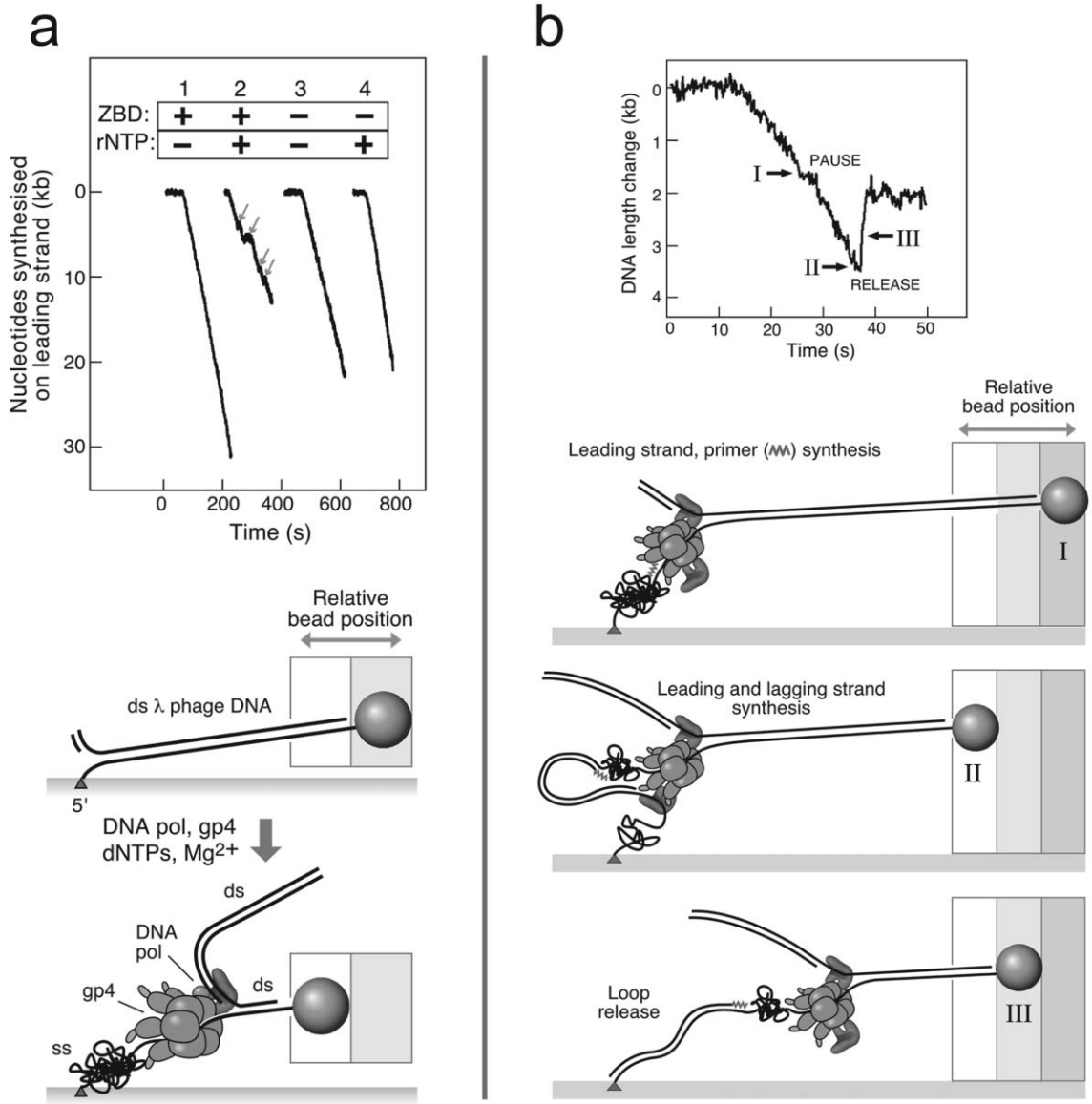

Fig. 6 Single-molecule observation of DNA replication. (a) Single-molecule trajectories of leading-strand synthesis (top panel). Above the traces is indicated whether ribonucleotides (rNTP) were present in the reaction mixture, and whether the gp4 contained the primase zinc-binding domain (ZBD), without which no primase activity can occur. The grey arrows denote pauses. The bottom panel shows a schematic depiction of events observed in top panel. In the absence of lagging-strand synthesis, leading-strand synthesis causes the 5′ tail of the DNA to be converted to the single-stranded form. Attachment of the 5′ end to the surface allows the monitoring of this conversion as a change in total length of the DNA. (b) Single-molecule trajectory of replication loop formation. In the presence of excess T7 DNA polymerase, lagging-strand synthesis is initiated after primer synthesis (indicated by the pause) and a replication loop is formed in the lagging strand. Replication loop release is clearly visible as instantaneous lengthening of DNA. The lower panels provide a schematic explanation of these events (Figure adapted from ref. 35).

defective in primase activity (Fig. 6a, traces 3 and 4). The observation that leading-strand synthesis momentarily stops during primer synthesis explains how the slow enzymatic events on the lagging strand take place without leading-strand synthesis progressing too far ahead of the lagging-strand synthesis.

The authors extended these experiments to include lagging-strand synthesis by introducing excess T7 DNA polymerase. The ensuing lagging-strand synthesis leads to the formation and release of a replication loop on the lagging strand, expressed in the single-molecule traces as a shortening of the DNA that is followed by an instantaneous lengthening (Fig. 6b). The pauses prior to the initiation of replication loop formation represent a transient halting of the whole replication fork during primer synthesis and delivery.

The replisome is faced with significant topological challenges when replicating DNA. Two identical copies of a polymerase are moving in parallel at the fork while synthesizing DNA in an anti-parallel fashion. One polymerase synthesizes in a continuous manner, the other repeats a finely tuned sequence of primer utilization, synthesis, and recycling. These single-molecule experiments demonstrate how the production of the two strands is coupled and how the three enzymatic activities (polymerization, unwinding, priming) are

coordinated. All organisms have evolved similar mechanisms to overcome the asymmetry at the fork. It is to be expected that these techniques will be directly applicable to other *in vitro* reconstituted replication systems (T4,[86] *E. Coli*,[87] SV40),[88] leading to a deeper understanding of the mechanisms underlying DNA replication.

Conclusion

During recent years, the application of single-molecule techniques to unravel the kinetics of complex enzymatic systems and to understand the underlying mechanisms has been expanded to the study of larger multi-enzyme systems. These efforts bring closer to realization the goal of understanding the complex networks of interactions between partners within the intricate molecular machineries that support various cellular processes. To date, these experiments focused on the activities of systems *in vitro*, reconstituted in nonphysiological environments. An important future direction will be the study of more complex enzymatic processes, *e.g.* eukaryotic DNA replication, at the single-molecule level in cell-free extracts, an environment that closely mimics the cellular context, but is still compatible with *in vitro* single-molecule techniques.

References

1 C. M. Coppin, J. T. Finer, J. A. Spudich and R. D. Vale, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 1913.
2 S. C. Kuo and M. P. Sheetz, *Science*, 1993, **260**, 232.
3 C. Kural, H. Kim, S. Syed, G. Goshima, V. I. Gelfand and P. R. Selvin, *Science*, 2005, **308**, 1469.
4 K. Svoboda, C. F. Schmidt, B. J. Schnapp and S. M. Block, *Nature*, 1993, **365**, 721.
5 K. Visscher, M. J. Schnitzer and S. M. Block, *Nature*, 1999, **400**, 184.
6 A. Yildiz, M. Tomishige, R. D. Vale and P. R. Selvin, *Science*, 2004, **303**, 676.
7 T. Funatsu, Y. Harada, M. Tokunaga, K. Saito and T. Yanagida, *Nature*, 1995, **374**, 555.
8 A. Yildiz, J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman and P. R. Selvin, *Science*, 2003, **300**, 2061.
9 C. Veigel, L. M. Coluccio, J. D. Jontes, J. C. Sparrow, R. A. Milligan and J. E. Molloy, *Nature*, 1999, **398**, 530.
10 J. E. Molloy, J. E. Burns, J. Kendrick-Jones, R. T. Tregear and D. C. White, *Nature*, 1995, **378**, 209.
11 J. T. Finer, R. M. Simmons and J. A. Spudich, *Nature*, 1994, **368**, 113.
12 B. Hille, 'Ion Channels of Excitable Membranes', Sinauer Associates, 2001.
13 R. A. Capaldi and R. Aggeler, *Trends Biochem. Sci.*, 2002, **27**, 154.
14 R. Yasuda, H. Noji, M. Yoshida, K. Kinosita, Jr. and H. Itoh, *Nature*, 2001, **410**, 898.
15 G. Kaim, M. Prummer, B. Sick, G. Zumofen, A. Renn, U. P. Wild and P. Dimroth, *FEBS Letters*, 2002, **525**, 156.
16 H. P. Lu, L. Xun and X. S. Xie, *Science*, 1998, **282**, 1877.
17 C. Bustamante, J. C. Macosko and G. J. L. Wuite, *Nature Rev. Mol.r Cell Biol.*, 2000, **1**, 130.
18 K. M. Herbert, A. La Porta, B. J. Wong, R. A. Mooney, K. C. Neuman, R. Landick and S. M. Block, *Cell*, 2006, **125**, 1083.
19 E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick and S. M. Block, *Nature*, 2005, **438**, 460.
20 J. W. Shaevitz, E. A. Abbondanzieri, R. Landick and S. M. Block, *Nature*, 2003, **426**, 684.
21 K. C. Neuman, E. A. Abbondanzieri, R. Landick, J. Gelles and S. M. Block, *Cell*, 2003, **115**, 437.
22 H. Yin, M. D. Wang, K. Svoboda, R. Landick, S. M. Block and J. Gelles, *Science*, 1995, **270**, 1653.
23 D. A. Schafer, J. Gelles, M. P. Sheetz and R. Landick, *Nature*, 1991, **352**, 444.
24 C. Joo, S. A. McKinney, M. Nakamura, I. Rasnik, S. Myong and T. Ha, *Cell*, 2006, **126**, 515.
25 J. van Mameren, M. Modesti, R. Kanaar, C. Wyman, G. J. Wuite and E. J. Peterman, *Biophys. J.*, 2006.
26 A. Graneli, C. C. Yeykal, R. B. Robertson and E. C. Greene, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 1221.
27 M. Spies, P. R. Bianco, M. S. Dillingham, N. Handa, R. J. Baskin and S. C. Kowalczykowski, *Cell*, 2003, **114**, 647.
28 N. Handa, P. R. Bianco, R. J. Baskin and S. C. Kowalczykowski, *Mol. Cell*, 2005, **17**, 745.
29 I. Amitani, R. J. Baskin and S. C. Kowalczykowski, *Mol. Cell*, 2006, **23**, 143.
30 P. R. Bianco, L. R. Brewer, M. Corzett, R. Balhorn, Y. Yeh, S. C. Kowalczykowski and R. J. Baskin, *Nature*, 2001, **409**, 374.
31 P. C. Blainey, A. M. van Oijen, A. Banerjee, G. L. Verdine and X. S. Xie, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 5752.
32 Z. Zhang, M. M. Spiering, M. A. Trakselis, F. T. Ishmael, J. Xi, S. J. Benkovic and G. G. Hammes, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 3254.
33 S. Myong, I. Rasnik, C. Joo, T. M. Lohman and T. Ha, *Nature*, 2005, **437**, 1321.
34 K. Pant, R. L. Karpel, I. Rouzina and M. C. Williams, *J. Mol. Biol.*, 2004, **336**, 851.
35 J.-B. Lee, R. K. Hite, S. M. Hamdan, X. S. Xie, C. C. Richardson and A. M. van Oijen, *Nature*, 2006, **439**, 621.
36 B. Maier, D. Bensimon and V. Croquette, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 12002.
37 G. J. L. Wuite, S. B. Smith, M. Young, D. Keller and C. Bustamante, *Nature*, 2000, **404**, 103.
38 S. Waga and B. Stillman, *Annu. Rev. Biochem.*, 1998, **67**, 721.
39 S. J. Benkovic, A. M. Valentine and F. Salinas, *Annu. Rev. Biochem.*, 2001, **70**, 181.
40 M. Chandler, R. E. Bird and L. Caro, *J. Mol. Biol.*, 1975, **94**, 127.
41 J. W. Drake, E. F. Allen, S. A. Forsberg, R. M. Preparata and E. O. Greening, *Nature*, 1969, **221**, 1128.
42 C. Bustamante, J. F. Marko, E. D. Siggia and S. Smith, *Science*, 1994, **265**, 1599.
43 J. F. Marko and E. D. Siggia, *Macromolecules*, 1994, **27**, 981.
44 A. M. van Oijen, P. C. Blainey, D. J. Crampton, C. C. Richardson, T. Ellenberger and X. S. Xie, *Science*, 2003, **301**, 1235.
45 T. R. Strick, J. F. Allemand, D. Bensimon, A. Bensimon and V. Croquette, *Science*, 1996, **271**, 1835.
46 T. R. Strick, V. Croquette and D. Bensimon, *Nature*, 2000, **404**, 901.
47 T. R. Strick, J. F. Allemand, D. Bensimon and V. Croquette, *Biophys. J.*, 1998, **74**, 2016.
48 J. Gore, Z. Bryant, M. D. Stone, M. Nollmann, N. R. Cozzarelli and C. Bustamante, *Nature*, 2006, **439**, 100.
49 A. Ashkin and J. M. Dziedzic, *Science*, 1987, **235**, 1517.
50 A. Ashkin, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 4853.
51 C. Bustamante, Y. R. Chemla, N. R. Forde and D. Izhaky, *Annu. Rev. Biochem.*, 2004, **73**, 705.
52 W. E. Moerner and D. P. Fromm, *Rev. Sci. Instrum.*, 2003, **74**, 3597.
53 D. Axelrod, *Met. Cell. Biol.*, 1989, **30**, 245.
54 P. R. Selvin, *Nat. Struct. Biol.*, 2000, **7**, 730.
55 A. A. Deniz, M. Dahan, J. R. Grunwell, T. Ha, A. E. Faulhaber, D. S. Chemla, S. Weiss and P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 3670.
56 S. A. McKinney, A. C. Declais, D. M. Lilley and T. Ha, *Nat. Struct. Biol.*, 2003, **10**, 93.
57 I. Rasnik, S. Myong, W. Cheng, T. M. Lohman and T. Ha, *J. Mol. Biol.*, 2004, **336**, 395.
58 A. Kornberg and T. A. Baker, *DNA Replication*, W. H. Freeman and Company, 1992.
59 I. Wong, S. S. Patel and K. A. Johnson, *Biochemistry*, 1991, **30**, 526.
60 S. Doublie, S. Tabor, A. M. Long, C. C. Richardson and T. Ellenberger, *Nature*, 1998, **391**, 251.
61 H. Yang, G. Luo, P. Karnchanaphanurach, T. M. Louie, I. Rech, S. Cova, L. Xun and X. S. Xie, *Science*, 2003, **302**, 262.
62 A. Goel, M. D. Frank-Kamenetskii, T. Ellenberger and D. Herschbach, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 8485.

63 A. Goel, R. D. Astumian and D. Herschbach, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 9699.
64 L. E. Bird, H. S. Subramanya and D. B. Wigley, *Curr. Opin. Struct. Biol.*, 1998, **8**, 14.
65 E. C. Friedberg, G. C. Walker, W. Siede, R. D. Wood, R. A. Schultz and T. Ellenberger, *DNA Repair and Mutagenesis*, ASM Press, 2006.
66 M. N. Dessinges, T. Lionnet, X. G. Xi, D. Bensimon and V. Croquette, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 6439.
67 K. M. Dohoney and J. Gelles, *Nature*, 2001, **409**, 370.
68 A. K. Eggleston and S. C. West, *Curr. Biol.*, 1997, **7**, R745.
69 T. T. Perkins, H. W. Li, R. V. Dalal, J. Gelles and S. M. Block, *Biophys. J.*, 2004, **86**, 1640.
70 T. Ha, I. Rasnik, W. Cheng, H. P. Babcock, G. H. Gauss, T. M. Lohman and S. Chu, *Nature*, 2002, **419**, 638.
71 I. Rasnik, S. Myong, W. Cheng, T. M. Lohman and T. Ha, *J. Mol. Biol.*, 2004, **336**, 395.
72 S. S. Patel and K. M. Picha, *Annu. Rev. Biochem.*, 2000, **69**, 651.
73 M. D. Betterton and F. Julicher, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2005, **71**, 011904.
74 S. C. Kowalczykowski, *Trends Biochem. Sci.*, 2000, **25**, 156.
75 K. Pant, R. L. Karpel and M. C. Williams, *J. Mol. Biol.*, 2003, **327**, 571.
76 M. A. Trakselis and S. J. Benkovic, *Structure*, 2001, **9**, 999.
77 R. D. Smiley, Z. Zhuang, S. J. Benkovic and G. G. Hammes, *Biochemistry*, 2006, **45**, 7990.
78 J. Xi, Z. Zhang, Z. Zhuang, J. Yang, M. M. Spiering, G. G. Hammes and S. J. Benkovic, *Biochemistry*, 2005, **44**, 7747.
79 J. Xi, Z. Zhuang, Z. Zhang, T. Selzer, M. M. Spiering, G. G. Hammes and S. J. Benkovic, *Biochemistry*, 2005, **44**, 2305.
80 F. Salinas and S. J. Benkovic, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 7196.
81 K. Tougu and K. J. Marians, *J. Biol. Chem.*, 1996, **271**, 21398.
82 J. Lee, P. D. Chastain, II, J. D. Griffith and C. C. Richardson, *J. Mol. Biol.*, 2002, **316**, 19.
83 S. Tabor, H. E. Huber and C. C. Richardson, *J. Biol. Chem.*, 1987, **262**, 16212.
84 D. N. Frick, K. Baradaran and C. C. Richardson, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 7957.
85 S. Guo, S. Tabor and C. C. Richardson, *J. Biol. Chem.*, 1999, **274**, 30303.
86 B. M. Alberts, J. Barry, P. Bedinger, T. Formosa, C. V. Jongeneel and K. N. Kreuzer, *Cold Spring Harb Symp Quant Biol*, 1983, **47 Pt 2**, 655.
87 C. A. Wu, E. L. Zechner and K. J. Marians, *J. Biol. Chem.*, 1992, **267**, 4030.
88 S. Waga and B. Stillman, *Nature*, 1994, **369**, 207.